JNMS Journal of Nippon Medical School https://www.nms.ac.jp/sh/jnms/

Original

Differences in Exhibited Emotions between Junior Residents and Senior Doctors: Using an Artificial Intelligence-Based Imaging Analysis Tool

Naoto Matsuda[1,2], Sonoko Kirinoki[1,2] and Gen Takagi[1]

[1]Department of General Medicine and Health Science, Nippon Medical School, Tokyo, Japan
[2]Department of General Medicine, Dokkyo Medical University Saitama Medical Center, Saitama, Japan

Background: Students and junior residents often rely on subjective methods and evaluations to learn medical interviewing. Although facial expressions correlate with therapeutic outcomes, no study has systematically analyzed facial expressions in medical education. This study aimed to investigate the differences in facial expressions between senior physicians and junior residents during medical interviews using an artificial intelligence (AI)-based facial expression analysis tool.

Methods: Healthcare professionals at the Dokkyo Medical University Saitama Medical Center were recruited between November 2017 and October 2018. The medical interview duration was compared between junior and senior physicians. Facial emotions were analyzed using "Kokoro Sensor," an AI-based tool that classifies facial expressions into seven emotions—anger, contempt, disgust, fear, joy, sadness, and surprise—based on the proportion of frames classified as each emotion.

Result: Thirteen physicians participated, resulting in 20 video recordings from 8 junior residents and 19 from 5 senior physicians. The mean interview time was 18.95 ± 8.959 minutes for junior residents and 11.79 ± 8.073 minutes for senior physicians ($p = 0.004$). The percentage of time physicians' faces were recognized by the Kokoro Sensor (indicating when doctors looked at patients) was 9.4% for junior residents and 21.6% for senior physicians ($p = 0.012$). Emotional analysis showed a significant difference in the expression of surprise: 4.9% for junior residents and 15.9% for senior physicians ($p = 0.011$).

Conclusion: Facial emotion analysis revealed differences in facial expressions between junior residents and senior physicians during medical interviews. The findings suggest that an AI-based facial expression analysis tool can be used in education for both students and residents.

(J Nippon Med Sch 2026; 93 (2): 190–196. https://doi.org/10.1272/jnms.JNMS.2026_93-213)

Keywords: facial emotion recognition, artificial intelligence, facial expression

Introduction

Recent advancements in information technology have led to the development of image-based evaluation software that is widely applied across multiple fields, including consumer behavior analysis and training programs aimed at improving communication skills[1,2]. Initially developed in the United States, facial expression recognition software has evolved rapidly with advances in computer processing power and video capture technology. Current applications include adaptive gaming systems responsive to user affect[3], tools for assessing viewer engagement with commercial media[4], and programs for enhancing nonverbal communication skills in professional training[5].

However, in the medical field, the integration of video informatics, including artificial intelligence (AI), remains relatively limited despite its growing potential. A limited but increasing body of literature demonstrates that facial

Correspondence to Naoto Matsuda, naoto-matsuda@nms.ac.jp
https://doi.org/10.1272/jnms.JNMS.2026_93-213
Received: October 17, 2025; Accepted: January 26, 2026

expression patterns may serve as clinically relevant indicators of mental health. For example, early studies using facial electromyography found that facial activity during initial clinical interviews could predict treatment outcomes in patients with depression[6]. More recent investigations using both manual and automated analysis techniques have shown that positive facial expressions, such as smiles, increase throughout a successful treatment of depression[7]. Similarly, in populations with eating disorders, those who recovered exhibited significantly more positive facial expressions than those still experiencing symptoms[8]. Notably, no studies have systematically analyzed facial expressions in the context of medical education, highlighting a critical gap in the literature.

Facial expressions play a crucial role in nonverbal communication during medical encounters. Mast[9] have demonstrated that the frequency of physicians' smiles is positively correlated with patient satisfaction, highlighting the impact of subtle facial cues on the clinical experience. Furthermore, a meta-analysis by Zolnierek and DiMatteo[10] found that physicians' communication skills, which include nonverbal elements such as facial expressions, are significantly associated with improved patient adherence to treatment. In a recent study involving simulated patients, Schneider et al.[11] reported that the presence of Duchenne smiles by medical students was associated with higher ratings of communication quality and patient comfort. Taken together, prior research highlights the clinical and educational significance of facial expressions, yet no studies have systematically analyzed physicians' facial expressions during real medical interviews. This represents a notable gap in the literature. We hypothesized that physicians' facial expressions during medical interviews would differ according to their level of clinical experience, with more experienced physicians demonstrating distinct patterns of emotional expression compared with junior residents.

This study investigated differences between facial expressions exhibited by junior residents and senior physicians during medical interviews. The following key parameters were analyzed: mean interview duration, patient-facing time, facial valence, and emotion-based expression profiles.

Materials and Methods

Participants and Study Period

Senior physicians and resident physicians in the Department of General Medicine, Dokkyo Medical University Saitama Medical Center, who provided consent, were recruited between November 2017 and October 2018. Only first-visit patients who provided informed consent were included. No disease-specific criteria were used. Written consent was obtained from the staff and resident physicians after the study was explained to them by the research team. Patient consent was obtained in written form after an explanation by the senior physicians.

This study was approved by the Ethics Committee of Dokkyo Medical University Koshigaya Hospital (Approval No. 1741) and by the Ethics Committee of Nippon Medical School (Approval No. A-2020-031) in accordance with the principles of the Declaration of Helsinki.

The principal investigator obtained renewed ethical approval (Approval No. A-2020-031) with the aim of expanding participant recruitment. However, the onset of the COVID-19 pandemic shortly thereafter prevented additional data collection. Consequently, all participants included in this study were recruited exclusively from Dokkyo Medical University Saitama Medical Center, ensuring that no mixed institutional settings influenced the study results.

Video Recording and Editing

The patients were equipped with a neck-mounted iPod touch holder for an iPod touch model A1367 for video recording. Recordings started before the beginning of the medical interview and ended after the consultation. Subsequently, Wondershare version 8.1.0 was used for editing to include the time from the start of the interview to the start of the physical examination.

Interview Duration

The interview duration was defined as the time from the start of the interview until the start of the physical examination. This period included the time spent typing on a keyboard and engaging in conversations unrelated to the patient's condition.

Emotion Recognition AI

Kokoro Sensor (Version 1.1.0.0) is an AI-based facial expression analysis tool that uses deep learning algorithms informed by the Facial Action Coding System (FACS), a widely accepted anatomical framework for identifying facial muscle movements. FACS, originally developed by Ekman and Friesen, provides a standardized method for coding facial action units that correspond to specific muscle activations and has remained the foundational structure for most contemporary facial expression recognition technologies[12]. Recent reviews have emphasized that although modern AI models have advanced in computational efficiency and training data availability, they continue to rely on FACS-derived representations to classify facial expressions and interpret emotional states[13,14]. In this study, seven core emotions were quantified using FACS-based action units detected by the Kokoro Sensor, enabling reproducible and anatomically grounded emotional analysis.

Seven emotions were identified using Kokoro Sensor. The system operates on a PC equipped with a camera and automatically generates emotion values after image analysis, requiring no specialized expertise. Each emotion—sadness, happiness, anger, surprise, fear, disgust, and contempt—was detected based on deep learning algorithms incorporating FACS and assigned numerical values ranging from 0 to 100 for each video frame. In this study, the presence of emotions was considered for values of 50 or higher.

For emotion analysis, the proportion of each emotion was calculated by dividing the number of frames with a value ≥ 50 for each emotion by the total number of frames with emotion values ≥ 50. Emotional valence was quantified on a continuous scale ranging from −100 to +100. For categorical analysis, the valence values were classified into the following five levels:

Strong positive: valence $\geq +80$
Mild positive: $+20 <$ valence $< +80$
Neutral: $-20 \leq$ valence $\leq +20$
Mild negative: $-80 <$ valence < -20
Strong negative: valence ≤ -80

The percentage of each valence category was calculated by dividing the number of frames in that category by the total number of frames for facial recognition.

Patient-Facing Time

Patient-facing time was defined as the percentage of video frames in which the physician's face was detected and analyzed by the Kokoro Sensor. It was set to analyze facial expressions only when the physician's face was oriented toward the patient. Therefore, the number of frames in which the facial expression was analyzed was divided by the total number of frames during the medical interview to calculate the percentage of time spent facing the patient. This percentage was defined as the ratio of the physician's facial orientation toward the patient.

Statistical Analysis

The Shapiro–Wilk test was used to assess the normality of the distributions. If normality was not confirmed, a non-parametric method, specifically the Mann–Whitney U test, was employed for group comparisons. Statistical significance was set at $p < 0.05$.

Results

Participants

Thirteen physicians participated in this study, comprising five senior physicians and eight junior residents, and 39 video recordings were analyzed. The experience of the senior physicians ranged from 3 to 11 years. The senior physicians included in this study had diverse but relevant clinical backgrounds. One male physician had 11 years of clinical experience and was board-certified in internal medicine and infectious diseases. He was affiliated with a university hospital and actively engaged in clinical practice, medical education, and research. Another female senior physician with 11 years of experience was board-certified in internal medicine, cardiology, and diabetology. She worked at a university hospital with responsibilities in education, research, and clinical care. A third senior physician had 3 years of post-graduate experience and was primarily dedicated to outpatient clinical practice, including the supervision and training of junior residents. Despite variations in their specialty backgrounds, all senior physicians routinely conducted medical interviews in general medicine settings, ensuring the comparability of their communication skills for the purposes of this study. All junior residents were in their first year of clinical training. All sessions were conducted during initial medical interviews in the Department of General Medicine.

Interview Duration

The mean interview duration was significantly longer for junior residents (18.95 ± 8.95 minutes) than for senior physicians (11.79 ± 8.07 minutes) ($p < 0.05$; **Figure 1**).

Patient-Facing Time

The analysis revealed that junior residents had an average patient-facing time of 9.4% (3,543 frames), whereas senior physicians' was 21.6% (4,009 frames) ($p < 0.05$; **Figure 2**).

Comparison of Expressiveness and Valence

Senior physicians demonstrated lower overall facial expressivity and a predominantly neutral emotional valence, whereas junior residents displayed a wider range of emotional expressions, both positive and negative (**Figure 3**).

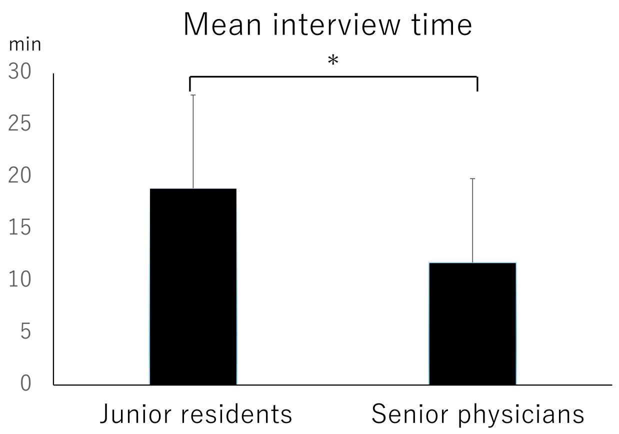

Figure 1 Mean interview duration
The mean interview duration was significantly longer for junior residents (18.95 ± 8.95 minutes) than for senior physicians (11.79 ± 8.07 minutes).
* p = 0.04 (p<0.05).

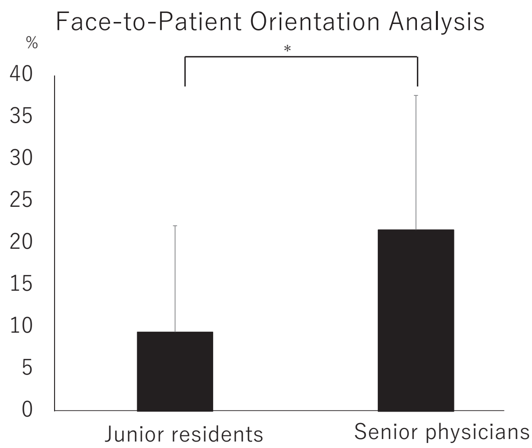

Figure 2 Patient-facing time analysis
The analysis of patient-facing time revealed that junior residents had an average patient-facing time of 9.4% (3,543 frames), whereas senior physicians' was 21.6% (4,009 frames).
* p = 0.012 (p<0.05).

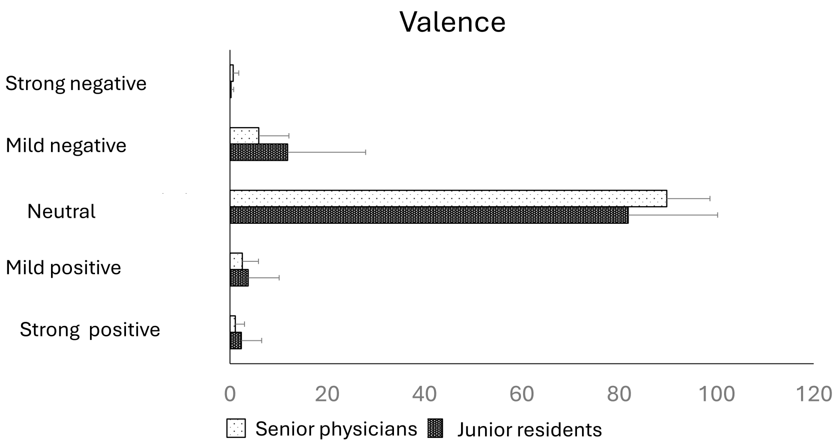

Figure 3 Comparison of expressiveness and valence
There is no significant difference between the groups.

Emotion Analysis

Regarding individual emotions, junior residents showed more joy, fear, and disgust, whereas senior physicians displayed surprise, sadness, and contempt more frequently. Anger was rarely observed in either group. A statistically significant difference was found in the expression of "surprise," which was notably more frequent among senior physicians (4.93 ± 16.0% in junior residents and 15.9 ± 26.9% in senior physicians) ($p < 0.05$; **Figure 4**).

Discussion

Facial expressions are increasingly recognized as relevant indicators of clinical diagnosis and patient interaction assessments. Previous studies have shown that patients with conditions such as anorexia nervosa, autistic spectrum disorder, depression, and Parkinson's disease exhibit distinct facial expression patterns that correlate with disease characteristics and progression[7,8,15,16].

In this study, the medical interview duration was notably longer among junior residents than it was among senior physicians. These findings indicate that experienced clinicians obtain patient information more efficiently through targeted questioning and well-developed communication skills, reflecting their advanced interviewing competence. Similarly, the higher patient-facing time observed among senior physicians likely reflects more patient-centered communication patterns, consistent with prior literature on expert–novice differences in clinical interviewing[17]. However, the lower patient-facing time

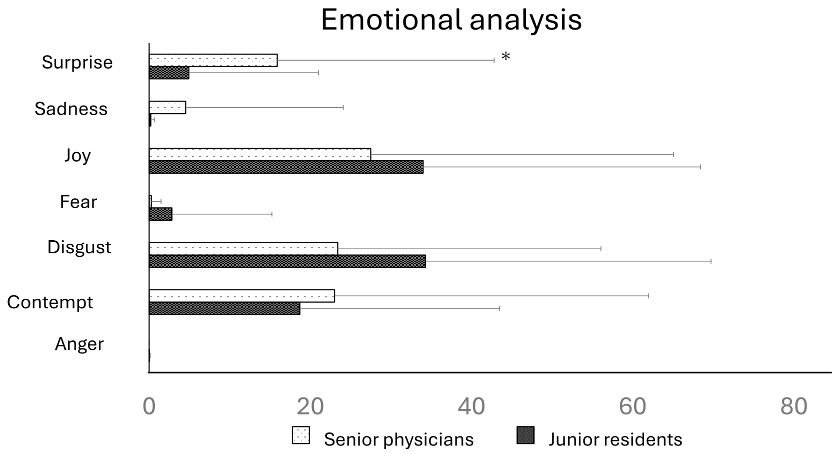

Figure 4 Emotion analysis
A statistically significant difference was found in the expression of "surprise," which was notably more frequent among senior physicians (4.93 ± 16.0% in junior residents and 15.9 ± 26.9% in senior physicians).
* p = 0.011 (p<0.05).

among junior residents also warrants consideration. Several factors may contribute to this finding. First, junior residents often require more time to operate the electronic medical record (EMR) system during consultations, which may reduce the amount of time they are able to face the patient directly. Second, the cognitive load associated with formulating diagnostic hypotheses, planning questions, and simultaneously documenting information may require frequent shifts of attention away from the patient. In contrast, senior physicians may perform EMR documentation more efficiently, allowing them to maintain visual engagement with patients for longer periods. While these explanations cannot be confirmed from our dataset alone, they represent plausible contributors and highlight the importance of integrating communication training with EMR-use skills in early post-graduate education.

Interestingly, although valence analysis did not yield significant differences between the groups, emotional analysis revealed a significant variation in the frequency of "surprise" expressions. Senior physicians displayed the expression of surprise more frequently than junior residents, despite our initial assumption that less experienced clinicians would show stronger emotional reactions. The higher frequency of surprise expressions among senior physicians should be interpreted with caution. Because facial expressions are influenced by multiple contextual, interpersonal, and situational factors, the underlying reason for this difference cannot be determined from our data. Rather than reflecting specific cognitive processes, such as the recognition of unexpected patient responses, this finding may represent natural variability in facial expressiveness across individuals and clinical encounters. Future studies incorporating qualitative analysis or synchronized contextual annotation would be required to clarify the potential meaning of these expressions. Alternatively, this finding may be an artifact of specific scenarios in the medical interviews recorded in this study.

Facial expression analysis software can objectively quantify facial expressions, which are key components of nonverbal communication. In Japanese medical education, the teaching of nonverbal communication has traditionally relied on experiential learning, leading to variability in instruction across institutions. This inconsistency has been highlighted as a limitation in objective structured clinical examinations[18]. By enabling the quantification of nonverbal communication behaviors, such technologies provide a pathway toward more objective and standardized assessments[19].

This advancement is expected to shift the current paradigm of medical education, which used to depend heavily on subjective experiences. This study demonstrated the potential of AI-based software for facial expression recognition, such as the Kokoro Sensor, to objectively quantify emotional responses and communication behaviors. These tools can provide novel methods for evaluating clinical communication skills, particularly in the context of training medical students and residents.

Limitations

This study has limitations. First, the relatively small sample size of the physicians and recorded medical interviews may have restricted the generalizability of the findings. Thus, the results may not be representative of all clinical settings, specialties, or levels of physician experience.

Second, all the medical interviews were conducted in the Department of General Medicine and involved only first-visit patients. This departmental and clinical context may influence communication dynamics and limit the applicability of the findings to other medical specialties or follow-up settings.

Third, while the use of an AI-based facial expression recognition system allows for objective analysis, the accuracy of such systems can vary depending on factors such as lighting conditions, camera angle, and individual facial characteristics. Additionally, AI does not consider contextual or cultural nuances in emotional expression, which may result in the misclassification or underestimation of certain affective states. Further validation with larger and more diverse samples and improved AI models is warranted.

Fourth, the study data were collected in 2017–2018, using an earlier version of the emotion-recognition tool. Although both physician groups were analyzed using the same system, changes in clinical communication practices, workflow environments, and technological development over time may limit the generalizability of the findings to current clinical settings.

Fifth, the interview recordings were not fully independent, as several interviews were conducted by the same physician. This clustering may introduce within-physician correlation, which is not accounted for by the Mann–Whitney U test used in our analysis. As a result, the p-values may be underestimated, leading to overestimation of statistical significance. Future studies should apply statistical methods that appropriately address repeated measures or hierarchical data structures, such as mixed-effects models, generalized estimating equations, or cluster-robust variance estimators.

Conclusion

This study employed a video-based analysis of physicians' facial orientation and facial expressions during medical interviews. The findings highlight a discrepancy in patient-facing time between senior and junior physicians, with the former showing greater eye contact and conducting more efficient interviews. AI-based systems for recognizing facial expressions have proven to be a viable tool for quantifying differences in facial expressions between junior and senior physicians. The findings suggest that AI-based technologies can be integrated into medical education to objectively assess communication skills and facilitate personalized training. However, further validation in a broader clinical setting is required to confirm our findings.

Author Contributions: Sonoko Kirinoki and Naoto Matsuda were responsible for patient recruitment and case registration. Naoto Matsuda drafted the entire manuscript, and Gen Takagi supervised the overall study.

Acknowledgments: The authors gratefully acknowledge Dr. Hajime Tamura for valuable contributions to patient recruitment and case registration.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The software used in this study was provided by CAC Corporation, and other study-related expenses were covered by departmental funds.

Conflict of Interest: The authors declare that they have no conflicts of interest related to this study.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process: During the preparation of this manuscript, the authors used ChatGPT (OpenAI) to assist with language editing and refinement of English expressions. After using this tool, the authors reviewed and edited the content as necessary and take full responsibility for the final version of the manuscript.

References

1. Zhang C, Wang X, Zhao C, et al. PromotionLens: Inspecting promotion strategies of online e-commerce via visual analytics. IEEE Trans Vis Comput Graph. 2023 Jan;29(1): 767–77.
2. Liciotti D, Contigiani M, Frontoni E, Mancini A, Zingaretti P, Placidi V. Shopper analytics: A customer activity recognition system using a distributed RGB-D camera network. Video Analytics for Audience Measurement: First International Workshop, VAAM 2014; 2014 Aug 24; Stockholm, Sweden: Springer; 2014 Nov 18.
3. Chowanda A, Dennis V, Dharmawan V, Ramli J. Player's affective states as Meta AI design on augmented reality games. Int J Inform Vis. 2023 Jun;7(2):561–8.
4. Mehta A, Sharma C, Kanala M, Thakur M, Harrison R, Torrico DD. Self-reported emotions and facial expressions on consumer acceptability: A study using energy drinks. Foods. 2021 Feb 4;10(2):330.
5. Jin E, Kang H, Lee K, Lee SG, Lee EC. Analysis of nursing students' nonverbal communication patterns during simulation practice: A pilot study. Healthcare (Basel). 2023 Aug 18;11(16):2335.

6. Greden JF, Price HL, Genero N, Feinberg M, Levine S. Facial EMG activity levels predict treatment outcome in depression. Psychiatry Res. 1984 Dec;13(4):345–52.
7. Girard JM, Cohn JF, Mahoor MH, Mavadati S, Rosenwald DP. Social risk and depression: Evidence from manual and automatic facial expression analysis. Proc Int Conf Autom Face Gesture Recognit. 2013:1–8.
8. Leppanen J, Dapelo MM, Davies H, Lang K, Treasure J, Tchanturia K. Computerised analysis of facial emotion expression in eating disorders. PLoS One. 2017 Jun 2;12(6): e0178972.
9. Mast MS. On the importance of nonverbal communication in the physician-patient interaction. Patient Educ Couns [Internet]. 2007 Aug;67(3):315–8. Available from: https://www.sciencedirect.com/science/article/pii/S0738399107000973?via%3Dihub. doi: 10.1016/j.pec.2007.03.005
10. Zolnierek KB, DiMatteo MR. Physician communication and patient adherence to treatment: A meta-analysis. Med Care [Internet]. 2009 Aug;47(8):826–34. Available from: https://journals.lww.com/lww-medicalcare/fulltext/2009/08000/physician_communication_and_patient_adherence_to.2.aspx. doi: 10.1097/MLR.0b013e31819a5acc
11. Schneider P, Zerbini G, Reicherts P, et al. Smiling doctor, satisfied patient-the impact of facial expressions on doctor-patient interactions. Front Med (Lausanne) [Internet]. 2025 Apr 24;12:1518517. Available from: https://www.frontiersin.org/journals/medicine/articles/10.3389/fmed.2025.1518517/full. doi: 10.3389/fmed.2025.1518517
12. Ekman P, Friesen WV. Facial action coding system: A technique for the measurement of facial movement. Palo Alto (CA): Consulting Psychologists Press; 1978.
13. Cohn JF, De la Torre F. Automated face analysis for affective computing. In: Calvo RA, D'Mello S, Gratch J, Kappas A, editors. The Oxford handbook of affective computing. Oxford: Oxford University Press; 2014. p. 131–50.
14. Martinez B, Valstar MF, Jiang B, Pantic M. Automatic analysis of facial actions: A survey. IEEE Trans Affective Comput. 2019 Jul-Sep 1;10(3):325–47.
15. Jin B, Qu Y, Zhang L, Gao Z. Diagnosing Parkinson disease through facial expression recognition: Video analysis. J Med Internet Res. 2020 Jul 10;22(7):e18697.
16. McIntosh DN, Reichmann-Decker A, Winkielman P, Wilbarger JL. When the social mirror breaks: Deficits in spontaneous, but not voluntary, mimicry of emotional facial expressions in autism. Dev Sci. 2006 May;9(3):295–302.
17. Yamada R, Xu K, Kondo S, Fujimoto M. Why the gaze behavior of expert physicians and novice medical students differ during a simulated medical interview: A mixed methods study. PLoS One. 2025 Jan 2;20(1): e0315405.
18. Ishikawa H, Hashimoto H, Kinoshita M, Fujimori S, Shimizu T, Yano E. Evaluating medical students' nonverbal communication during the objective structured clinical examination. Med Educ [Internet]. 2006 Dec;40(12):1180–7. Available from: https://asmepublications.onlinelibrary.wiley.com/doi/10.1111/j.1365-2929.2006.02628.x. doi: 10.1111/j.1365-2929.2006.02628.x
19. Yu IC, Guo JM, Lin WC, Fang JT. Development of nonverbal communication behavior model for nursing students based on deep learning facial expression recognition technology. Cogent Educ [Internet]. 2025;12(1): 2448059. Available from: https://www.researchgate.net/publication/387678839_Development_of_nonverbal_communication_behavior_model_for_nursing_students_based_on_deep_learning_facial_expression_recognition_technology. doi: 10.1080/2331186X.2024.2448059